TECHNISCHE UNIVERSITÄT DRESDEN
Fakultät Wirtschaftswissenschaften

Dresdner Beiträge zu
Quantitativen Verfahren

Nr. 38/04

Simultane Validierung von Ausfallwahrscheinlichkeiten

von

Andreas Henking

Herausgeber:
Die Professoren der
Fachgruppe Quantitative Verfahren

ISSN 0945-4802

Simultane Validierung von Ausfallwahrscheinlichkeiten

Andreas Henking

www.risk-sim.de

21. Mai 2004

1 Einleitung

In den Basler Konsultationspapieren wird empfohlen, die Ausfallwahrscheinlichkeiten der einzelnen Ratingklassen eines Ratings mit den beobachten Ausfallquoten ex post zu überprüfen (vgl. Basel Committee on Banking Supervision (2003), Absatz 464). Dies führt zu dem statistischen Standardtest auf Anteilswerte, der in einführender statistischer Literatur beschrieben ist (vgl. Fahrmeir et al. (2001)). Eine wichtige Voraussetzung für diesen Test ist, dass die Ausfälle aller Kreditnehmer, die zur Überprüfung der Ausfallwahrscheinlichkeiten herangezogen werden, stochastisch unabhängig sind. Diese Forderung dürfte in der Regel nicht erfüllt sein. Höse, Huschens (2003) leiten her, dass bei Vorliegen stochastischer Abhängigkeiten der Test nicht geeignet ist. Allerdings gehen die Autoren von einer recht hohen Assetkorrelation von 0,2 aus. Vor allem für Retailportfolien kann eine deutlich kleinere Assetkorrelation angenommen werden. Tasche (2003) zeigt, dass der Test, legt man für die Ausfallquoten eine Verteilung unter positiver Assetkorrelation zugrunde, später zur Ablehnung führt. Umgekehrt bedeutet dies, dass der Test unter fälschlicherweise angenommener Unabhängigkeit zu früh die Nullhypothese ablehnt. Hingegen halten Blochwitz et al. (2003) den Fehler unter Nichtberücksichtigung von Abhängigkeiten bei gewissen Parameterkonstellationen für vernachlässigbar und entwickeln ein mehrstufiges Testverfahren auf Basis des Standardtests.

Ein weiteres Problem bei der Überprüfung der Ausfallwahrscheinlichkeiten ist, dass gemäß Empfehlung der Basler Konsultationspapiere für jede Ratingklasse ein gesonderter Test durchgeführt wird. So sind, falls die Nullhypothese erfüllt ist, beim gleichzeitigen Testen der Ausfallwahrscheinlichkeiten von 12 Ratingklassen bei einem Signifikanzniveau von 10% Ablehnungen der Tests in 1 oder 2 Ratingklassen durchaus zu erwarten. Es muss also eine Vorschrift gefunden werden, die festlegt, ab wie vielen signifikanten Ergebnissen die Ausfallwahrscheinlichkeiten des Ratingsystems auf den Prüfstand gestellt und die Ausfallwahrscheinlichkeiten neu kalibriert werden müssen. Henking et al. (2004) schlagen hierzu, unter der Unabhängigkeitsannahme, ein zweistufiges Warnsystem bei 12 Ratingklassen vor.

Bei der Auswahl der Teststrategie ist auch die Motivation für das Backtesting zu beachten. Bei einem Backtesting im aufsichtsrechtlichen Sinne sollte der Test nicht zu früh auf eine Neukalibrierung hindeuten, um den Entwicklungsaufwand in Grenzen zu halten. Bei einem Backtesting im Sinne der internen Überwachung des Ratings, also zur Erzeugung von Frühwarnindikatoren zur Validität des Ratings, kann es durchaus erwünscht sein, dass der Test frühzeitig anschlägt, um weitere Untersuchungen einzuleiten (vgl. Henking et al.(2004)).

Aus obigen Gründen erscheint es notwendig, das Ablehnverhalten des Tests, also die Fehler 1. und 2. Art, unter verschiedenen Konstellationen zu untersuchen. Grundlage hierfür ist ein Ein-Faktor-Modell, über welches die Assetkorrelation gesteuert werden kann. Zwei Varianten des Modells, mit abhängigen und unabhängigen Ratingklassen, werden in Abschnitt 2 vorgestellt und in einer Simulationsstudie untersucht. Der Test mit Testgrößen und zugehörigen Hypothesen wird in Abschnitt 3 beschrieben. Zur Untersuchung des Fehlers 2. Art werden die Ausfallwahrscheinlichkeiten mit verschiedenen Faktoren erhöht. Der simultane Test auf allen Ratingklassen wird exemplarisch für 12 Ratingklassen untersucht. Dabei wird die Anzahl der beobachteten Perioden zwischen 1 und 5 variiert. Die Simulationsergebnisse sind in Abschnitt 4 aufgeführt.

2 Die Simulationsmodelle

Die Frage, die im folgenden beantwortet werden soll, ist, wie sich der Test verhält, wenn die Annahme der Unabhängigkeit verletzt ist. Dazu wird das Ein-Faktor-Modell (vgl. Höse, Huschens (2003), Bluhm et al. (2003)) für mehrere Perioden verwendet und die Fehler 1. und 2. Art werden über Simulationen bestimmt.

Auf Basis der Modelle werden jeweils Portfolien der Größe 5.000 in 10.000 Durchläufen über eine und über fünf Perioden simuliert. Die zwölf Ratingklassen teilen sich nach folgendem Schlüssel auf die Portfolien auf, mit den zugehörigen prognostizierten Ausfallwahrscheinlichkeiten (PD: Probability of Default):

Bereich	Anteil in %	PD in %
A	20	0.85
B	20	1.69
C	10	2.46
D	10	3.13
E	10	4.16
F	10	5.51
G	10	8.03
H	5	12.17
I	2	16.21
K	1	19.76

Bereich	Anteil in %	PD in %
L	1	23.96
M	1	34.75

Tabelle 1: Verteilung auf die Ratingklassen und die Ausfallwahrscheinlichkeiten der Ratingklassen in den Simulationsportfolien

Die Aufteilung und die zugehörigen Ausfallwahrscheinlichkeiten in Tabelle 1 entsprechen denen des SCHUFA-Score für Banken (vgl. www.schufa.de). Zu beachten ist, dass die Ratingklassen K, L und M in den simulierten Portfolien jeweils nur 50 Kreditnehmer enthalten.

2.1 Simulationsmodell I

Im Ein-Faktor-Modell tritt der Ausfall D_{ij}^t des j-ten Kreditnehmers in Ratingklasse i in Periode t ein, wenn dessen standardisierte Bonitätsänderung r_{ij}^t unter dem Schwellenwert c_{ij} liegt:

$$D_{ij}^t = \mathbf{1}\left\{r_{ij}^t < c_{ij}\right\},\ i = 1,\ldots,m\ ,\ j = 1,\ldots,n_i\ ,\ t = 1,\ldots,T$$

$$r_{ij}^t = \sqrt{\rho} \cdot Y_t + \sqrt{1-\rho} \cdot \varepsilon_{ij}^t$$

Dabei bezeichnet **1** die Indikatorfunktion, Y_t sind die stochastisch unabhängigen standardnormal verteilten *systematischen* Faktoren, ε_{ij}^t sind die untereinander und von den Y_t stochastisch unabhängigen standardnormal verteilten *spezifischen* Faktoren, ρ ist die *Assetkorrelation*. Die Ausfallwahrscheinlichkeiten der einzelnen Ratingklassen werden über die Perioden als konstant angenommen. Damit gilt folgendes:

$$\begin{aligned} r_{ij}^t &\sim N(0,1) \\ p_i &= P\left(r_{ij}^t < c_{ij}\right) \\ c_{ij} &= \Phi^{-1}\left(p_i\right). \end{aligned}$$

Die systematischen Faktoren Y_t sind unabhängig von den Ratingklassen und innerhalb einer Periode für alle Kreditnehmer gleich. Die Ausfallwahrscheinlichkeit p_i ist über alle Perioden t konstant.

2.2 Simulationsmodell II

Im Simulationsmodell I wird in der Abhängigkeitsstruktur nicht zwischen den Ratingklassen unterschieden. Dies führt zu Simulationsmodell II, in welchem die Ausfälle der Kreditnehmer verschiedener Ratingklassen als stochastisch unabhängig angenommen werden. Damit stellt sich das Modell wie folgt dar:

$$r_{ij}^t = \sqrt{\rho_i} \cdot Y_{t,i} + \sqrt{1-\rho_i} \cdot \varepsilon_{ij}^t$$

wobei die $Y_{t,i}$ als stochastisch unabhängig angenommen werden. Ansonsten ergeben sich keine Unterschiede zwischen den Modellen. Die Assetkorrelation ist somit spezifisch für jede Ratingklasse. Bei den Simulationsberechnungen jedoch wird aus Vereinfachungsgründen in der Darstellung von der gleichen Assetkorrelation in allen Ratingklassen ausgegangen. Daher kann in der Modellgleichung die Indizierung der Assetkorrelation ρ fallen gelassen werden.

3 Beschreibung des Tests

Als Testgröße dienen die beobachteten Ausfallquoten $\hat{p}_i$ je Ratingklasse i. Die Anzahl der Ratingklassen ist prinzipiell beliebig. Die beobachtete Ausfallquote $\hat{p}_i$, als Schätzer für die Ausfallwahrscheinlichkeit p_i in Ratingklasse i, ist die Anzahl der Ausfälle einer Ratingklasse im Verhältnis zur Gesamtzahl n_i der Kreditnehmer dieser Ratingklasse. Werden T Perioden betrachtet, dann bezeichnet $\hat{p}_{i,t}$ die beobachtete Ausfallquote in Ratingklasse i in Periode t, $n_{i,t}$ die Portfoliogröße von Ratingklasse i in Periode t und es ist

$$n_i = \sum_{t=1}^{T} n_{i,t}$$
$$\hat{p}_i = \frac{1}{n_i} \sum_{t=1}^{T} n_{i,t} \cdot \hat{p}_{i,t}.$$

Sind die Ausfälle der Kreditnehmer stochastisch unabhängig, so ist $\hat{p}_i$ binomial verteilt. Näherungsweise ist $\hat{p}_i$ normal verteilt. Eine Faustregel besagt, dass für $n_i \cdot p_i \cdot (1-p_i) > 9$ die Näherung durch die Normalverteilung ausreichend gut ist. Wir testen einseitig die Hypothese

$$p_i = p_i^0 \text{ gegen } p_i > p_i^0.$$

Die Nullhypothese ist unter der Unabhängigkeitsannahme abzulehnen, falls

$$\hat{p}_i > p_i^0 + z_\alpha \sqrt{\frac{p_i^0 \left(1-p_i^0\right)}{n_i}} =: p_{i,krit},$$

wobei α das Signifikanzniveau dieses Test und z_α das α - Quantil der Standardnormalverteilung ist. Der Fehler 1. Art, also die fälschlicherweise Ablehnung der Nullhypothese, hat die Wahrscheinlichkeit α. Der Fehler 2. Art, die fälschlicherweise Beibehaltung der Nullhypothese, hängt von der Lage der wahren Ausfallwahrscheinlichkeit p_i ab.

Bei dem simultanen Test der Ausfallwahrscheinlichkeiten aller m Ratingklassen wird die Hypothese

$$p_i = p_i^0 \text{ für } i = 1,\dots,m \text{ gegen } p_i > p_i^0 \text{ für mindestens ein } i \in \{1,\dots,m\}$$

überprüft.

Seien für alle Ratingklassen die Unabhängigkeitsannahme und die Nullhypothese erfüllt, so erwartet man bei gleichzeitiger Durchführung des Tests für alle k Ratingklassen, dass für $k \cdot \alpha$ Ratingklassen der Test trotzdem zur Ablehnung führt. Für $k = 12$ Ratingklassen und einem Signifikanzniveau von 10% bedeutet dies, dass durchschnittlich in 1,2 Ratingklassen der Test zur Ablehnung führt. Anders ausgedrückt: mit einer Wahrscheinlichkeit von 0,72 führt in mindestens einer Ratingklasse der Test zur Ablehnung. Bei einem Signifikanzniveau von 0,05 beträgt diese Wahrscheinlichkeit immerhin noch 0,46. Daher wird vorgeschlagen (vgl. Henking et al. (2004)) für die nächst größere ganze Zahl von $k \cdot \alpha$ die Warnstufe auf *gelb* zu setzen. Warnstufe *gelb* ist ein Hinweis darauf, die Ausfallwahrscheinlichkeiten und die Ursachen für die Abweichung genauer zu untersuchen. Die Definition einer Warnstufe *rot* ist schwieriger in eine allgemeine Faustregel zu fassen, sollte aber deutlich jenseits der Warnstufe gelb liegen und von fachlichen Überlegungen geleitet sein.

Die Faustregel für die Näherung durch die Normalverteilung kann je nach Portfoliogröße z.B. in den Ratingklasse K, L und M verletzt sein (vgl. Tabelle 1). Dies ist jedoch bewusst so gewählt, da davon auszugehen ist, dass in der Anwendung mit dem Normalverteilungsquantil gearbeitet wird, ohne der Einhaltung der Faustregel besonderes Augenmerk zu schenken.

4 Ergebnisse

Es werden für verschiedene Signifikanzniveaus (Spalte $1-\alpha$) und verschiedene Assetkorrelationen (Spalte ρ) die mittlere Ablehnhäufigkeiten des Tests ermittelt. Die Ergebnisse werden für Ratingklasse A und Ratingklasse M, sowie für das gleichzeitige Testen aller 12 Ratingklassen angegeben.

4.1 Simulationsmodell I

4.1.1 Fehler 1. Art

In Tabelle 2 sind die Ergebnisse für eine einjährige Betrachtungsweise angegeben, in Tabelle 3 jene bei einer fünfjährigen Betrachtungsweise. Den Ergebnissen liegen die Ausfallwahrscheinlichkeiten aus Tabelle 1 zugrunde, die zugleich auch die Nullhypothesen darstellen. Die mittleren Ablehnhäufigkeiten für die Ratingklassen A und M entsprechen somit dem durch die Simulation abgeschätzten Fehler 1. Art. Die mittleren Ablehnhäufigkeiten aller 12 Ratingklassen können zur Konstruktion einer Testvorschrift für den simultanen Test verwendet werden, bei vorgegebenen Signifikanzniveau und in

Abhängigkeit von der Assetkorrelation. Allerdings muss der kritische Wert auf ganze Zahlen gerundet werden, was zu einem anderen Signifikanzniveau führt.

		Mittlere Ablehnhäufigkeit		
$1-\alpha$	ρ	**Ratingklasse A**	**Ratingklasse M**	**Alle 12 Ratingklassen**
0,99	0	0,012	0,0089	0,15
0,99	0,01	0,044	0,019	0,56
0,99	0,02	0,074	0,029	0,92
0,99	0,05	0,12	0,067	1,6
0,99	0,1	0,15	0,12	2,1
0,99	0,2	0,15	0,19	2,4
0,95	0	0,048	0,064	0,68
0,95	0,01	0,099	0,091	1,4
0,95	0,02	0,13	0,11	1,8
0,95	0,05	0,17	0,17	2,4
0,95	0,1	0,19	0,22	2,7
0,95	0,2	0,19	0,26	2,9
0,90	0	0,089	0,11	1,3
0,90	0,01	0,14	0,14	2,0
0,90	0,02	0,17	0,17	2,4
0,90	0,05	0,20	0,22	2,9
0,90	0,1	0,22	0,27	3,1
0,90	0,2	0,20	0,31	3,2

Tabelle 2: Ablehnhäufigkeiten im einperiodischen Simulationsmodell I, falls Nullhypothese erfüllt (Fehler 1. Art)

Bereits für die kleine Assetkorrelation $\rho = 0{,}01$ resultieren deutliche Unterschiede im Fehler 1. Art gegenüber dem Unabhängigkeitsfall, also $\rho = 0$.

		Mittlere Ablehnhäufigkeit		
$1-\alpha$	ρ	**Ratingklasse A**	**Ratingklasse M**	**Alle 12 Ratingklassen**
0,99	0	0,013	0,0099	0,15
0,99	0,01	0,047	0,021	0,52
0,99	0,02	0,077	0,033	0,89
0,99	0,05	0,14	0,069	1,6
0,99	0,1	0,19	0,12	2,3
0,99	0,2	0,23	0,19	2,9
0,95	0	0,048	0,044	0,60
0,95	0,01	0,10	0,066	1,3
0,95	0,02	0,13	0,093	1,7
0,95	0,05	0,20	0,14	2,5
0,95	0,1	0,24	0,20	3,0

0,95	0,2	0,26	0,26	3,5
0,90	0	0,11	0,098	1,2
0,90	0,01	0,17	0,13	2,0
0,90	0,02	0,20	0,16	2,4
0,90	0,05	0,25	0,21	3,1
0,90	0,1	0,28	0,26	3,5
0,90	0,2	0,29	0,30	3,9

Tabelle 3: Ablehnhäufigkeiten im fünfperiodischen Simulationsmodell I, falls Nullhypothese erfüllt (Fehler 1. Art)

Der Fehler 1. Art ist unabhängig von der Anzahl der Perioden und der Portfoliogröße, wie dies dem allgemeinen Konstruktionsprinzip statistischer Tests entspricht. Der Fehler 1. Art hängt somit nur vom Signifikanzniveau und der Assetkorrelation ab.

4.1.2 Fehler 2. Art

Zur Einschätzung des Fehlers 2. Art wurden im Simulationsmodell die mit den Faktoren 1,1 (Tabelle 4 und Tabelle 6) und 1,2 (Tabelle 5 und Tabelle 7) multiplizierten Ausfallwahrscheinlichkeiten verwendet. Getestet wurde gegen die angenommenen Ausfallwahrscheinlichkeiten in Tabelle 1. Für die Ratingklassen A und M ist die Schätzung für den Fehler 2. Art: $1 - \textit{mittlere Ablehnhäufigkeit}$. Das bedeutet, je höher die mittleren Ablehnhäufigkeiten, desto geringer ist der Fehler 2. Art.

		Mittlere Ablehnhäufigkeit		
$1-\alpha$	ρ	**Ratingklasse A**	**Ratingklasse M**	**Alle 12 Ratingklassen**
0,99	0	0,027	0,032	0,46
0,99	0,01	0,074	0,051	1,1
0,99	0,02	0,10	0,071	1,5
0,99	0,05	0,15	0,13	2,1
0,99	0,1	0,18	0,18	2,6
0,99	0,2	0,17	0,25	2,8
0,95	0	0,090	0,16	1,6
0,95	0,01	0,15	0,19	2,3
0,95	0,02	0,17	0,22	2,7
0,95	0,05	0,21	0,26	3,1
0,95	0,1	0,22	0,31	3,4
0,95	0,2	0,21	0,34	3,4
0,90	0	0,15	0,25	2,6
0,90	0,01	0,20	0,27	3,2
0,90	0,02	0,22	0,29	3,4
0,90	0,05	0,25	0,32	3,7

0,90	0,1	0,25	0,36	3,8
0,90	0,2	0,22	0,38	3,7

Tabelle 4: Ablehnhäufigkeiten im einperiodischen Simulationsmodell I, falls Nullhypothese nicht erfüllt (Fehler 2. Art) mit um Faktor 1,1 erhöhten Ausfallwahrscheinlichkeiten

		Mittlere Ablehnhäufigkeit		
$1-\alpha$	ρ	**Ratingklasse A**	**Ratingklasse M**	**Alle 12 Ratingklassen**
0,99	0	0,055	0,093	1,1
0,99	0,01	0,11	0,12	1,9
0,99	0,02	0,14	0,15	2,3
0,99	0,05	0,19	0,20	2,8
0,99	0,1	0,20	0,26	3,1
0,99	0,2	0,20	0,31	3,3
0,95	0	0,15	0,32	3,0
0,95	0,01	0,20	0,34	3,5
0,95	0,02	0,23	0,35	3,7
0,95	0,05	0,25	0,38	4,0
0,95	0,1	0,25	0,40	4,0
0,95	0,2	0,23	0,42	3,9
0,90	0	0,23	0,42	4,3
0,90	0,01	0,26	0,43	4,5
0,90	0,02	0,28	0,44	4,6
0,90	0,05	0,29	0,45	4,6
0,90	0,1	0,28	0,45	4,5
0,90	0,2	0,25	0,46	4,2

Tabelle 5: Ablehnhäufigkeiten im einperiodischen Simulationsmodell I, falls Nullhypothese nicht erfüllt (Fehler 2. Art) um Faktor 1,2 erhöhten Ausfallwahrscheinlichkeiten

		Mittlere Ablehnhäufigkeit		
$1-\alpha$	ρ	**Ratingklasse A**	**Ratingklasse M**	**Alle 12 Ratingklassen**
0,99	0	0,061	0,12	1,3
0,99	0,01	0,12	0,15	2,0
0,99	0,02	0,15	0,18	2,5
0,99	0,05	0,22	0,23	3,1
0,99	0,1	0,26	0,28	3,5
0,99	0,2	0,28	0,33	3,9
0,95	0	0,16	0,30	3,2
0,95	0,01	0,21	0,32	3,7
0,95	0,02	0,25	0,34	4,0
0,95	0,05	0,29	0,36	4,3

0,95	0,1	0,31	0,39	4,5
0,95	0,2	0,31	0,42	4,6
0,90	0	0,28	0,45	4,8
0,90	0,01	0,32	0,45	4,9
0,90	0,02	0,33	0,45	5,0
0,90	0,05	0,36	0,46	5,1
0,90	0,1	0,36	0,46	5,1
0,90	0,2	0,34	0,47	5,0

Tabelle 6: Ablehnhäufigkeiten im fünfperiodischen Simulationsmodell I, falls Nullhypothese nicht erfüllt (Fehler 2. Art) mit um Faktor 1,1 erhöhten Ausfallwahrscheinlichkeiten

		Mittlere Ablehnhäufigkeit		
$1-\alpha$	ρ	**Ratingklasse A**	**Ratingklasse M**	**Alle 12 Ratingklassen**
0,99	0	0,18	0,49	4,7
0,99	0,01	0,24	0,48	4,8
0,99	0,02	0,26	0,48	4,9
0,99	0,05	0,31	0,48	5,0
0,99	0,1	0,33	0,48	5,0
0,99	0,2	0,32	0,49	4,9
0,95	0	0,35	0,73	7,4
0,95	0,01	0,37	0,70	6,9
0,95	0,02	0,38	0,68	6,7
0,95	0,05	0,39	0,64	6,3
0,95	0,1	0,39	0,61	6,0
0,95	0,2	0,36	0,58	5,6
0,90	0	0,51	0,84	8,8
0,90	0,01	0,50	0,81	8,1
0,90	0,02	0,49	0,78	7,7
0,90	0,05	0,47	0,73	7,1
0,90	0,1	0,44	0,68	6,6
0,90	0,2	0,40	0,63	6,1

Tabelle 7: Ablehnhäufigkeiten im fünfperiodischen Simulationsmodell I, falls Nullhypothese nicht erfüllt (Fehler 2. Art) mit um Faktor 1,2 erhöhten Ausfallwahrscheinlichkeiten

Die Fehler 2. Art im fünfperiodischen Simulationsmodell sind wesentlich geringer als im einperiodischen Modell. Was im wesentlich daran liegt, dass die Perioden als stochastisch unabhängig modelliert sind, womit eine Reduzierung der Varianz des Schätzers $\hat{p}_i$ einher geht. Damit ist die Empfehlung zu geben, mehrere Perioden in die Analyse einzubeziehen, wie auch von Höse, Huschens (2003) vorgeschlagen. Zudem bietet sich bei der Betrachtung über mehrere Perioden auch die Möglichkeit, Trends in den Ausfallwahrscheinlichkeiten zu betrachten.

Tabelle 7 liefert das interessante und wichtige Ergebnis, dass der Fehler 2. Art nicht monoton steigend mit der Assetkorrelation sein muss. Dies ist teilweise auch in den vorhergehenden Simulationsergebnissen beim Vergleich der Ergebnisse zwischen $\rho = 0{,}1$ und $\rho = 0{,}2$ zu beobachten.

4.2 Simulationsmodell I vs Simulationsmodell II

Die Tests einzelner Ratingklassen sind für die Simulationsmodelle I und II äquivalent. Unterschiede können sich also nur für $\rho \neq 0$ beim gleichzeitigen Testen mehrerer Ratingklassen ergeben. Es resultieren auch hier für die beiden Modelle die gleichen Fehler 1. und 2. Art. Jedoch verringert sich die Standardabweichung der Ablehnhäufigkeiten für das Simulationsmodell II im Bereich der Alternativhypothese recht deutlich.

Unter der Nullhypothese bei Unabhängigkeit, also $\rho = 0$, lässt sich die Standardabweichung der Ablehnhäufigkeiten, unter der Annahme dass die Normalapproximation des Tests hinreichend genau ist, theoretisch bestimmen. Sie hängt nur von dem Signifikanzniveau und der Anzahl der Ratingklassen ab. Insbesondere ist sie damit unabhängig vom Simulationsmodell. Für 12 Ratingklassen sind die Standardabweichungen in Abhängigkeit vom Signifikanzniveau in Tabelle 8 angegeben.

1-Signifikanzniveau	Standardabweichung
0,99	0,34
0,95	0,75
0,90	1,04

Tabelle 8: Standardabweichungen der Ablehnhäufigkeiten unter der Nullhypothese für 12 Ratingklassen in Abhängigkeit vom Signifikanzniveau und für $\rho = 0$

Für den Bereich der Alternativhypothese sind die Standardabweichungen der Ablehnhäufigkeiten in Tabelle 9 exemplarisch für ein Signifikanzniveau von 0,05 und eine Assetkorrelation von 0,02 angegeben. Dabei zeigt die Spalte „Faktor PD“ mit welchem Faktor die Ausfallwahrscheinlichkeiten in der Simulation multipliziert wurden. Ein Faktor = 1 entspricht der Nullhypothese, ein Faktor > 1 der Alternativhypothese.

		Standardabweichung der Ablehnhäufigkeit	
Anzahl Perioden	Faktor PD	Modell I	Modell II
1	1	2,0	1,2
1	1,1	4,0	1,4
5	1	1,9	1,2
5	1,1	4,4	1,6

Tabelle 9: Vergleich der Standardabweichungen der Ablehnhäufigkeiten für die beiden Simulationsmodelle im Bereich der Alternativhypothese, für $\alpha = 0{,}05$ und $\rho = 0{,}02$.

Man erkennt eine deutliche Reduktion der Standardabweichung im Simulationsmodell II. Wenn die Annahmen des Simulationsmodells II plausibel erscheinen, ist somit eine exaktere Warnstufendefinition als im Simulationsmodell I möglich.

4.3 Ergebnisse für andere Portfoliogrößen im Simulationsmodell I

Es stellt sich die Frage, inwieweit die Portfoliogröße das Ablehnverhalten bei den Tests steuert. Dazu wurde die Simulationsstudie mit zwei weiteren Portfoliogrößen, 1.000 und 20.000, durchgeführt. Bei der kleineren Portfoliogröße mit 1.000 Kreditnehmern besteht zudem die Gefahr, dass die Normalapproximation bei der Bestimmung des kritischen Wertes nicht mehr so gut greift. In diesem Fall finden sich in Ratingklasse M beispielsweise nur noch 10 Kreditnehmer.

Anhand der folgenden Simulationsergebnissen lässt sich die generelle Aussage aufstellen, dass mit größeren Portfolioumfang die Ablehnneigung der Tests steigt. Dies bedeutet bei kleineren Portfolioumfängen einen kleineren Fehler 1. Art und einen größeren Fehler 2. Art als bei größeren Portfolioumfängen.

4.3.1 Fehler 1. Art

		Mittlere Ablehnhäufigkeit		
$1-\alpha$	ρ	**Ratingklasse A**	**Ratingklasse M**	**Alle 12 Ratingklassen**
0,99	0	0,031	0,0066	0,23
0,99	0,05	0,073	0,045	0,70
0,95	0	0,092	0,090	0,78
0,95	0,05	0,13	0,12	1,4

Tabelle 10: Portfoliogröße 1.000; Ablehnhäufigkeiten im einperiodischen Simulationsmodell I, falls Nullhypothese erfüllt (Fehler 1. Art)

		Mittlere Ablehnhäufigkeit		
$1-\alpha$	ρ	**Ratingklasse A**	**Ratingklasse M**	**Alle 12 Ratingklassen**
0,99	0	0,012	0,009	0,14
0,99	0,05	0,21	0,19	2,8
0,95	0	0,054	0,052	0,65
0,95	0,05	0,26	0,27	3,4

Tabelle 11: Portfoliogröße 20.000; Ablehnhäufigkeiten im einperiodischen Simulationsmodell I, falls Nullhypothese erfüllt (Fehler 1. Art)

4.3.2 Fehler 2. Art

		Mittlere Ablehnhäufigkeit		
$1-\alpha$	ρ	**Ratingklasse A**	**Ratingklasse M**	**Alle 12 Ratingklassen**
0,99	0	0,043	0,043	0,38
0,99	0,05	0,089	0,067	0,95
0,95	0	0,12	0,14	1,2
0,95	0,05	0,16	0,17	1,8

Tabelle 12: Portfoliogröße 1.000; Ablehnhäufigkeiten im einperiodischen Simulationsmodell I, falls Nullhypothese nicht erfüllt (Fehler 2. Art) mit um Faktor 1,1 erhöhten Ausfallwahrscheinlichkeiten

		Mittlere Ablehnhäufigkeit		
$1-\alpha$	ρ	**Ratingklasse A**	**Ratingklasse M**	**Alle 12 Ratingklassen**
0,99	0	0,029	0,031	0,46
0,99	0,05	0,084	0,051	1,1
0,95	0	0,09	0,16	1,6
0,95	0,05	0,15	0,19	2,3

Tabelle 13: Portfoliogröße 1.000; Ablehnhäufigkeiten im fünfperiodischen Simulationsmodell I, falls Nullhypothese nicht erfüllt (Fehler 2. Art) mit um Faktor 1,1 erhöhten Ausfallwahrscheinlichkeiten

		Mittlere Ablehnhäufigkeit		
$1-\alpha$	ρ	**Ratingklasse A**	**Ratingklasse M**	**Alle 12 Ratingklassen**
0,99	0	0,054	0,093	1,1
0,99	0,05	0,26	0,31	3,7
0,95	0	0,16	0,28	3,0
0,95	0,05	0,31	0,40	4,4

Tabelle 14: Portfoliogröße 20.000; Ablehnhäufigkeiten im einperiodischen Simulationsmodell I, falls Nullhypothese nicht erfüllt (Fehler 2. Art) mit um Faktor 1,1 erhöhten Ausfallwahrscheinlichkeiten

		Mittlere Ablehnhäufigkeit		
$1-\alpha$	ρ	**Ratingklasse A**	**Ratingklasse M**	**Alle 12 Ratingklassen**
0,99	0	0,15	0,49	4,7
0,99	0,05	0,36	0,49	5,2
0,95	0	0,36	0,74	7,5
0,95	0,05	0,42	0,59	6,1

Tabelle 15: Portfoliogröße 20.000; Ablehnhäufigkeiten im fünfperiodischen Simulationsmodell I, falls Nullhypothese nicht erfüllt (Fehler 2. Art) mit um Faktor 1,1 erhöhten Ausfallwahrscheinlichkeiten

5 Schlussfolgerungen

Die Simulationsstudie dient dazu, die Auswirkungen auf den Fehler 1. und 2. Art bei der Validierung von Ausfallwahrscheinlichkeiten zu beleuchten, falls fälschlicherweise von der Unabhängigkeit der Kreditnehmer eines Portfolien ausgegangen wird. Zugleich können aus den Simulationsergebnissen Ablehnvorschriften für das simultane Testen aller Ratingklassen abgeleitet werden. Als Nebenergebnis kann die Güte der Anpassung an die Grenzverteilung des Ein-Faktor-Modells für verschiedene Parameterkonstellationen untersucht werden.

Bereits für die kleine Assetkorrelation 0,01 reagieren die Test in ihrem Ablehnverhalten deutlich. Daher sollten vorhandene Abhängigkeiten berücksichtigt werden, indem die Ablehnbereiche bzw. Warnstufen modifiziert werden. So sollte bei einem Signifikanzniveau von 0,05 unter der Kenntnis einer schwachen Assetkorrelation (≤ 0,02) die Warnstufe gelb beim gleichzeitigen Testen von 12 Ratingklassen erst gesetzt werden, wenn die Tests für mindestens zwei Ratingklassen signifikante Ergebnisse aufweisen. Dies gewährleistet die näherungsweise Einhaltung des Fehlers 1. Art.

Ein wichtiges Ergebnis ist, dass für den Fehler 1. Art ohne Bedeutung ist, ob die Ausfallquote über fünf oder nur eine Periode geschätzt wird. Der Fehler 2. Art hingegen wird bei der Betrachtung von fünf Perioden deutlich kleiner.

Bevor die Ausfallwahrscheinlichkeiten der einzelnen Ratingklassen getestet werden, sollte ein Test auf die mittlere Ausfallquote, also der Gesamtausfallquote des Portfolien, vorgeschaltet sein. Weicht die beobachtete Gesamtausfallquote stark von der angenommenen ab, so kann dies ein Hinweis auf vorhandene positive Asset- bzw. Ausfallkorrelation sein. Wird dies über mehrere Perioden beobachtet, ist das ein Hinweis darauf, dass sich die mittlere Ausfallquote nachhaltig geändert hat. Das Rating sollte auf diese Tatsache hin neu kalibriert werden.

Bei den Betrachtungen zum Fehler 2. Art ist zu beachten, dass die Ausfallwahrscheinlichkeiten alle mit dem gleichen Faktor in eine Richtung geändert wurden. Dies ist natürlich nur im Labor möglich und entspricht nicht unbedingt der Realität. Dennoch sind die Ergebnisse zum Fehler 2. Art recht zufrieden stellend, was aber zum Teil auf Kosten eines erhöhten Fehlers 1. Art geht. Es ist also auch möglich, die Warnstufendefinition nach dem Fehler zweiter Art auszurichten, da es wichtiger sein kann, Änderungen bei den Ausfallwahrscheinlichkeiten auf jeden Fall aufzudecken, in Kauf nehmend, dass auch fälschlicherweise Änderungen angezeigt werden.

6 Ausblick

Um die Gefahr der Risikounterschätzung zu untersuchen, wurde der einseitige Test untersucht. Von Interesse kann aber auch das Verhalten des Test für zweiseitige Fragestellungen sein, also

$$p_i = p_i^0 \text{ gegen } p_i \neq p_i^0 .$$

Hierbei sind jedoch ähnliche Ergebnisse, also erhöhte Ablehnquoten bei Abhängigkeit gegenüber Unabhängigkeit der Kreditnehmer, wie bei dem einseitigen Test zu erwarten.

Mit dem verwendeten Simulationsprogramm ist es möglich, den Test auch für Ratingsysteme mit einer anderen Anzahl Ratingklassen, anderen Ausfallwahrscheinlichkeiten und einer anderen Verteilung auf die Ratingklassen sowie für verschiedene Portfoliogrößen, zu untersuchen. Ebenso können die Alternativhypothesen verschieden gewählt werden. Es ist auch denkbar die Unabhängigkeitsannahme zwischen den Perioden fallen zu lassen, wie dies bei Tasche (2004) geschieht.

Anstelle des Tests unter Unabhängigkeitsannahme, kann auch die unter Modellannahmen „wahre" Verteilung der Ausfallquoten verwendet werden. So bildet Balthazar (2004) Konfidenzintervalle für die Ausfallwahrscheinlichkeit auf Basis der simulierten Verteilungen im Ein-Faktor-Modell endlicher Portfolien. Dies kann ebenfalls auf eine beliebige Anzahl von Ratingklassen verallgemeinert werden.

Der Autor dankt Helmut Küchenhoff, Leiter des statistischen Beratungslabors an der Universität München, für gewissenhaftes Korrekturlesen und wertvolle Hinweise.

Literatur

Basel Committee on Banking Supervision (2003): „Consultative Document: The New Basel Capital Accord". April 2003.

Balthazar, L. (2004): „PD estimates for Basel II". Risk April 2004.

Blochwitz, S., Wehn, C., Hohl, S. (2003): „Reconsidering Ratings". Working Paper. November 2003.

Bluhm, C., Overbeck, L., Wagner, C. (2003): „Credit Risk Modelling", Chapman & Hall.

Fahrmeir, L., Künstler, R., Pigeot, I., Tutz, G. (2001): „Statistik. Der Weg zur Datenanalyse", Springer Verlag.

Henking, A., Hüls, R., Krieger, S. (2004): „Monitoring und Backtesting von Retail-Scoresystemen”. Working Paper. April 2004.

Höse, S., Huschens, S. (2003): „Sind interne Ratingsysteme im Rahmen von Basel II evaluierbar? Zur Schätzung von Ausfallwahrscheinlichkeiten durch Ausfallquoten“. Zeitschrift für Betriebswirtschaft 73.

Tasche, D. (2003): „A traffic lights approach to PD validation”. `http://www.defaultrisk.com`.

Tasche, D. (2004): „Simulation study on the performance of the normal and traffic lights tests”. Working paper. Februar 2004.

Anhang

Für das einperiodische Simulationsmodell kann die Verteilungsfunktion der Grenzverteilung der Ausfallquote $\hat{p}_i$ angegeben werden (vgl. Höse, Huschens (2003)):

$$F_{\hat{p}_i}(x) = \Phi\left(\frac{\sqrt{1-\rho}\,\Phi^{-1}(x) - \Phi^{-1}(p_i)}{\sqrt{\rho}}\right).$$

Mit der Ablehnvorschrift des Tests kann nun die aus der Grenzverteilung abgeleitete theoretische Ablehnwahrscheinlichkeit ermittelt werden, die dem Wert der Verteilungsfunktion $1-F_{\hat{p}_i}(p_{i,krit})$ entspricht. Die zu Tabelle 2 korrespondierenden theoretischen Werte sind in Tabelle 16 angegeben. Zu beachten ist, dass über $p_{i,krit}$ der Umfang der jeweiligen Ratingklasse im Portfolio in die Berechnung eingeht.

		Theoretische Ablehnwahrscheinlichkeit	
$1-\alpha$	ρ	**Ratingklasse A**	**Ratingklasse M**
0,99	0	0,01	0,01
0,99	0,01	0,0096	0,000029
0,99	0,02	0,042	0,0022
0,99	0,05	0,11	0,036
0,99	0,1	0,14	0,10
0,95	0	0,05	0,05
0,95	0,01	0,036	0,002
0,95	0,02	0,089	0,021
0,95	0,05	0,16	0,097
0,95	0,1	0,19	0,18
0,90	0	0,10	0,10
0,90	0,01	0,070	0,012
0,90	0,02	0,13	0,054

0,90	0,05	0,19	0,15
0,90	0,1	0,21	0,23

Tabelle 16: Theoretische Ablehnhäufigkeiten im einperiodischen Simulationsmodell I, falls Nullhypothese erfüllt (Fehler 1. Art)

Beim Vergleich von Tabelle 2 und Tabelle 16 erkennt man, dass vor allem für kleine ρ die Grenzverteilung keine ausreichend genaue Annäherung liefert. Dies leuchtet unmittelbar ein, da für $\rho = 0$ die Verteilungsfunktion der Grenzverteilung nicht definiert ist. Für $\rho = 0$ ist jedoch die exakte Verteilung, bis auf die Näherung der kritischen Werte durch die Normalverteilung, bekannt, über die die zugehörigen Werte in Tabelle 16 bestimmt sind. Aus den Werten für größere ρ lässt sich erkennen, dass die Anzahl der gewählten simulierten Portfolien offenbar groß genug ist, um eine ausreichende Genauigkeit in den Ergebnissen zu erhalten.

In Tabelle 17 sind die zu Tabelle 4 zugehörigen theoretischen Ablehnhäufigkeiten angegeben. Für $\rho \neq 0$ basieren diese wieder, wie bei Tabelle 16, auf der Grenzverteilung für die relativen Häufigkeiten im Ein-Faktor-Modell. Es zeigt sich wieder, dass die Näherung durch die Grenzverteilung für kleine ρ nicht besonders gut ist.

		Theoretische Ablehnwahrscheinlichkeit	
$1-\alpha$	ρ	**Ratingklasse A**	**Ratingklasse M**
0,99	0	0,026	0,038
0,99	0,01	0,023	0,00097
0,99	0,02	0,069	0,014
0,99	0,05	0,14	0,083
0,99	0,1	0,17	0,16
0,95	0	0,098	0,13
0,95	0,01	0,074	0,025
0,95	0,02	0,14	0,083
0,95	0,05	0,20	0,19
0,95	0,1	0,22	0,26
0,90	0	0,17	0,23
0,90	0,01	0,13	0,091
0,90	0,02	0,19	0,17
0,90	0,05	0,24	0,27
0,90	0,1	0,25	0,33

Tabelle 17: Theoretische Ablehnhäufigkeiten im einperiodischen Simulationsmodell I, falls Nullhypothese nicht erfüllt (Fehler 2. Art) mit um Faktor 1,1 erhöhten Ausfallwahrscheinlichkeiten

Dresdner Beiträge zu Quantitativen Verfahren (ISSN 0945-4802)

Ältere Ausgaben (1994 – 7/97): www.tu-dresden.de/wwqvs/VaR/qvreihe.htm

8/97 **S. Huschens**: Genauigkeit von Schätzungen des Risikopotentials

9/97 **S. Huschens**: Confidence intervals for the Value-at-Risk

10/97 **S. Huschens**: Konfidenzintervalle für den Value-at-Risk

11/97 **S. Huschens**: Risikoabschätzung durch historische Simulation

12/97 **A. Henking**: Some approaches in order to model bivariate densities with fixed marginals

13/97 **R. Roth**: Die Bestimmung des At-the-money-Punktes europäischer Optionen - Implikationen für die Einführung neuer Basispreise an der DTB

14/97 **R. Roth**: Die Eignung eines Futures auf implizite Forwardvolatilitäten zum Handeln des Vega-Risikos von Optionen

15/97 **M. Brechtmann**: Wochentagseffekte am deutschen Aktienmarkt unter Berücksichtigung von ARCH-Effekten

16/98 **R. Roth**: Der VOLAX-Future - Ein Derivat zum Handeln des Vega-Risikos von Optionen

17/98 **S. Huschens**: Messung des besonderen Kursrisikos durch Varianzzerlegung

18/98 **S. Huschens** (Hrsg.): Value-at-Risk-Schlaglichter, Ausgabe 1/98

19/98 **S. Huschens**: Historische Simulation

20/98 **J. F. Kiviet, G. D. A. Phillips, B. Schipp**: Alternative bias approximations in first order reduced form models

21/98 **R. Roth**: Die Bewertung des VOLAX-Futures mit dem Arbitrageansatz

22/98 **H. W. Brachinger, U. Steinhauser**: Konzepte zur Messung von Risiko - vom intuitiven Risikobegriff zum Value at Risk

23/98 **S. Huschens** (Hrsg.): Value-at-Risk-Schlaglichter, Ausgabe 2/98

24/98 **S. Huschens, J.-R. Kim:** Measuring risk in Value-at-Risk based on Student's t-distribution

25/98 **S. Huschens, J.-R. Kim:** Measuring Risk in Value-at-Risk in the Presence of Infinite Variance

26/99 **S. Huschens, J.-R. Kim:** Blue for β in CAPM with Infinite Variance

27/99 **R. Prinzler:** Reliability of neural network-based Value-at-Risk estimates

28/99 **S. Huschens:** Anmerkungen zur Value-at-Risk-Definition

29/99 **S. Huschens:** Verfahren zur Value-at-Risk-Berechnung

30/00 **S. Huschens:** Value-at-Risk-Berechnung durch historische Simulation

31/00 **S. Huschens:** Von der Markt- zur Kreditrisikomessung

32/02 **S. Höse, S. Huschens:** Sind interne Ratingsysteme im Rahmen von Basel II evaluierbar? Zur Schätzung von Ausfallwahrscheinlichkeiten durch Ausfallquoten

33/03 **S. Höse, S. Huschens:** From Credit Scores to Stable Default Probabilities: A Model Based Approach

34/03 **S. Höse, S. Huschens:** Estimation of Default Probabilities in a Single-Factor Model

35/03 **S. Höse, S. Huschens:** Simultaneous Confidence Intervals for Default Probabilities

36/03 **S. Huschens, K. Vogl, R. Wania:** Estimation of Default Probabilities and Default Correlations

37/04 **K. Vogl, R. Wania:** BLUEs for Default Probabilities

38/04 **A. Henking:** Simultane Validierung von Ausfallwahrscheinlichkeiten